1645ALLEN

#11/7R 09/18/99

1 PAGE:

44

RAW SEQUENCE LISTING PATENT APPLICATION US/09/056,019A DATE: 07/07/1999 TIME: 14:28:17

Input Set: 1056019A.RAW

This Raw Listing contains the General Information Section and up to first 5 pages.

```
<110> APPLICANT: Tuomanen, Elaine I
                                             ENTERED
          Wizemann, Theresa
2
          Masure, H. R.
3
          Johnson, Leslie S.
4
    <120> TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
5
          CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
6
7
          THEREFROM AND USES THEREOF
8
    <130> FILE REFERENCE: 1340-1-017
9
    <140> CURRENT APPLICATION NUMBER: US/09/056,019A
10
    <141> CURRENT FILING DATE: 1998-04-07
11
     <160> NUMBER OF SEQ ID NOS: 39
12
     <170> SOFTWARE: PatentIn Ver. 2.0
13
     <210> SEQ ID NO 1
14
     <211> LENGTH: 406
15
     <212> TYPE: PRT
16
     <213> ORGANISM: Streptococcus pneumoniae
17
     <400> SEQUENCE: 1
           Glu Asn Glu Gly Ala Thr Gln Val Pro Thr Ser Ser Asn Arg Ala Asn
18
19
                                                 10
           Glu Ser Gln Ala Glu Gln Gly Glu Gln Pro Lys Lys Leu Asp Ser Glu
20
21
                                             25
           Arg Asp Lys Ala Arg Lys Glu Val Glu Glu Tyr Val Lys Lys Ile Val
22
23
           Gly Glu Ser Tyr Ala Lys Ser Thr Lys Lys Arg His Thr Ile Thr Val
24
25
                                     55
           Ala Leu Val Asn Glu Leu Asn Asn Ile Lys Asn Glu Tyr Leu Asn Lys
26
27
                                                     75
            Ile Val Glu Ser Thr Ser Glu Ser Gln Leu Gln Ile Leu Met Met Glu
28
 29
                                                 90
                             85
            Ser Arg Ser Lys Val Asp Glu Ala Val Ser Lys Phe Glu Lys Asp Ser
 30
 31
                                            105
                        100
            Ser Ser Ser Ser Ser Ser Ser Ser Thr Lys Pro Glu Ala Ser Asp
 32
 33
                                        120
            Thr Ala Lys Pro Asn Lys Pro Thr Glu Pro Gly Glu Lys Val Ala Glu
 34
 35
                                                         140
                                    135
            Ala Lys Lys Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys
 36
 37
                                                     155
                                 150
            Glu Glu Asp Arg Arg Asn Tyr Pro Thr Ile Thr Tyr Lys Thr Leu Glu
 38
 39
                                                 170
            Leu Glu Ile Ala Glu Ser Asp Val Glu Val Lys Lys Ala Glu Leu Glu
 40
 41
                                                                 190
                                             185
            Leu Val Lys Val Lys Ala Asn Glu Pro Arg Asp Glu Gln Lys Ile Lys
 42
 43
                                                             205
                                         200
                    195
```

DATE: 07/07/1999 RAW SEQUENCE LISTING 2 TIME: 14:28:17 PAGE:

PATENT APPLICATION US/09/056,019A

Input Set: 1056019A.RAW

45								215				Ala	Z Z U					
46			210	Tle	Lvs	Thr	Asp	Arq	Glu	Glu	Ala	Glu	Glu	Glu	Ala	Lys	Arg	
47							230					233						
48		225	λla	Δen	Δla	Lvs	Glu	Gln	Gly	Lys	Pro	Lys	Gly	Arg	Ala	Lys	Arg	
49						245					250					233		
50		Cl v	va 1	Pro	Glv	Glu	Leu	Ala	Thr	Pro	Asp	Lys	Lys	Glu	Asn	Asp	Ala	
51					200					265					2/0			
52 53		Tare	Ser	Ser	Asp	Ser	Ser	Val	Gly	Glu	Glu	Thr	Leu	$\mathtt{Pro}$	Ser	Pro	Ser	•
53 54				275					280					203				
54 55		Len	Lvs	Pro	Glu	Lys	Lys	Val	Ala	Glu	Ala	Glu	Lys	Lys	Val	Glu	GIu	L
56								295					300					
57		Δla	Lvs	Lvs	Lys	Ala	Glu	Asp	Gln	Lys	Glu	Glu	Asp	Arg	Arg	Asn	Tyr	:
58							210					313						
59		Pro	Thr	Asn	Thr	Tyr	Lys	Thr	Leu	Glu	Leu	Glu	Ile	Ala	GLu	ser	Asp	)
60						225					330							
61		Val	Glu	Val	Lys	Lys	Ala	Glu	Leu	Glu	Leu	Val	Lys	Glu	GIU	Ala	гу	3
62					240					345					330			
63		Glu	Pro	Arg	Asn	Glu	Glu	Lys	Val	Lys	Gln	Ala	Lys	Ala	GIU	. vai	GI	1
64				255					360					303				
65		Ser	Lys	Lys	Ala	Glu	Ala	Thr	Arg	Leu	Glu	Lys	11e	гуѕ	TIII	ASF	AL	3
66			270					375					300					
67		Lys	Lys	Ala	Glu	Glu	Glu	Ala	Lys	Arg	l PAs	Ala	Ата	Glu	GIU	. ASE	40°	0
68		385					390					395					10	•
69		Val	Lys	Glu	Lys	Pro	Ala											
70						405												
71	<210>																	
72	<211>				i													
73	<212>	TYP	E: P	RT														
74	<213>					toco	occus	pne	eumor	llae								
75	<400>	SEÇ	UENC	E: 2			<b>1</b>			ו די	ո Thi	r Gei	· Sei	^ Asr	Arc	r Ala	a As	n
76		Glu	ı Asr	Glu	ı GIŞ			GII	ı va.	L PI	111	r Ser				1	5	
77		1	_				. al-		. (1)	. (2)	n Dro	о Б	LVS	s Leu	ı Ası	o Se	r Gl	u
78		Glu	ı Ser	Glr			ı GII	1 GI	y GI	2	r F	J	, <u></u> , .		3	0		
79				_	20	) - 3	T	. (1)	, Wai	 1	u ഭിവ	u Tyi	r Vai	l Lvs			e Va	1
80		Arg	J Asi			a Arg	1 rA	S GI	u va. 41	n n	u 01	w - 1 -		45	5			
81		_		35	-		_ T		ri Th∵	0 12 T.37	s Tiv	s Arg	y Hi:			e Th	r Va	1
82		Gl			r Tyl	C AL	я груг	5 5E 5	E 111.	L LJ	5 27	J,	6	0				
83		_	50	) 	1 7	- 01	. To	ა 7\ლ ე	n Va.	n Tl	e Lv	s Ası	n Gl	u Ty	r Le	u As	n Ly	/S
84				ı va.	ı ASI	1 GI	7	n n	II AD.		· -1	7	5	-			ε	30
85		6!	5 			∽ Th	, GD:	r Gl	ıı Se	r Gl	n Le	u Gl	n Il	e Le	u Me	t Me	t G]	Lu
86						0	<b>=</b>				9	U				_	_	
87		<b>~</b> -	*\	~ 0~	~ T **	o e 17 e	_ ე გალ	n (31	u Al	a Va	l Se	r Ly	s Ph	e Gl	u Ly	s As	p Se	er
88					10	Λ				10	כ							
89		a -	- O-	~ 0~	r co	v r ge	r Se	r As	p Se	r Se	r Th	r Ly	s Pr	o Gl	u Al	a Se	r As	gz
90				11	_				12	0				14	5			
91		ml	~ 1 <sup>1</sup>	1.1. 1.1. c	o Dr	O As	n Lv	s Pr	o Th	r Gl	u Pr	o Gl	y Gl	u Ly	s Va	l Al	a G	lu
92			4.3	^				13	.5				14	U				
93		7. T	1.37 2. T.37	o Tar	g Tav	s Va	1 G1	u Gl	u Al	a Gl	.u Ly	s Ly	s Al	а Ly	s As	p Gl	n L	ys
94		AT	а шу	y سy	S Lly	_ • • •					-	_						

DATE: 07/07/1999 TIME: 14:28:17 RAW SEQUENCE LISTING
PATENT APPLICATION US/09/056,019A PAGE: 3

Input Set: 1056019A.RAW

						150					155					160
95	145		Asp .	7 ~~	λνα	130 Agn	Tvr	Pro	Thr	Ile	Thr	Tyr	Lys	Thr	Leu	Glu
96	GIU	Gru	Asp .		165	.1011	- 2 -			170		_			175	
97	T	<b>~1.</b> ,	Ile .	מות	Glu	Ser	Asp	Val	Glu	Val	Lys	Lys	Ala	Glu	Leu	Glu
98				100					185					TOO		
99	7	770 1	Lys	100 17a l	Taye	Δla	Asn	Glu	Pro	Arq	Asp	Glu	Gln	Lys	Ile	Lys
100			105					200					205			
101	a1	71-	Glu	פות	Glu	Va 1	Glu	Ser	Lvs	Gln	Ala	Glu	Ala	Thr	Arg	Leu
102		010					215					220				
103	_	210	Ile	T	Thr	λen	Ara	Glu	Glu	Ala	Glu	Glu	Glu	Ala	Lys	Arg
104		гуу	TTE	пур	1111	230	5				235					240
105	225		Asp	7 J -	Tare	230 2311	Gl n	Glv	Lvs	Pro	Lys	Gly	Arg	Ala	Lys	Arg
106	Arg	Ala	Asp	Ald	цуБ 245	Giu	G111	0-7	1-	250	-	-			255	
107			Pro	a1	243 2111	T.011	Δla	Thr	Pro	Asp	Lvs	Lys	Glu	Asn	Asp	Ala
108	GТĀ	Vai	Pro		GIU	пеи	AIG	1111	265			-		270		
109	_		Ser	260	Cox	Cor	172 I	Glv	Glu	Glu	Thr	Leu	Pro	Ser	Pro	Ser
110	Lys	ser		Asp	ser	ser	val	280	014	<b></b>			285			
111		_	275 Pro	a1	T	T	1751	Δla	G111	Ala	Glu	Lvs	Lys	Val	Glu	Glu
112	Leu		Pro	GIU	гуѕ	гуя	295	AIG	GIG	1114		300	•			
113	_	290	Lys	<b>. .</b>	31.	<b>~1.</b> ,	293	Gln.	T.VS	Glu	Glu		Arq	Arg	Asn	Tyr
114		Lys	Lys	гÀг	Ala		Asp	GIII	<b></b>	014	315	<b>L</b>	•	Ī		320
115	305		Asn	1		310	mb x	T.011	Glu	T.em		Ile	Ala	Glu	Ser	Asp
116	Pro	Thr	Asn	Thr		пув	1111	пец	GIU	330			_		335	
117	_	_	Val	_	325	71-	~1 <sub>11</sub>	Lou	Glu			Lvs	Glu	Glu	Ala	Lys
118	Val	Glu	Val		гуз	Ala	GIU	neu	345			-1-		350		_
119				340	<b>~</b> 1	<b>a</b> 1	T	77-7	Tare	G]n	Δla	Lvs	Ala	Glu	Val	Glu
120	Glu	Pro		Asn	GIU	GIU	пуъ	360	. Lys	. 011		-1-	365			
121			355		~1	314	mb ~	300	T.A11	c G1 11	Lvs	Ile			Asp	Arg
122	Ser			Ala	GIU	Ala	375	Arg	пес	. 014		380			• -	
123		370			<b>~</b> 3	~1	3/3	Taro	λνο	r Tays	ıΔla			Glu	Asp	Lys 400
124			Ala	GIU	GIU			шуз	, ,,,	, _, ~	395				_	400
125	385	; 	~-1	<b>.</b>	D	390	C111	Gln	Pro	Glr			Pro	Ala	Pro	Lys
126	Val	. Lys	GIU	гуѕ			. GIU	GIL		410	 }				415	1
127	_		_		405	Desc		Dro	. T.379			Asr	Pro	Ala	Glu	Gln
128	Ala	t GIU	Lys			PIC	ALC	L FIC	425	;				430	)	
129				420	•	Desc	. nl-	. Aer	- 22.	Glr	n Ala	Gli	ı Glü	. Asp	Tyr	Ala
130	Pro	Lys			гра	PIC	) Alc	440	)	. 0		-	445	5	_	
131		_	435		. a1	. al.	. The sec	770	, 2 2 2 0	r T.e.	ı Thi	Glr	ı Glr	ı Glr	n Pro	Pro
132	Arg			GIU	GIU	GIU	455	. <i>.</i>		,		460	)			
133		450	,	<b>.</b>		. 11-	. Clr	o Ora	5 GE	r Thi	r Pro	LVS	3 Thi	c Gly	/ Trp	Lys 480
134			GIU	гга	PIC	470	, GT:	1 11	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		475	5				480
135	46	5			. 1/0+	4 / L	) ` '''\!	r Dhe	<u></u> ጉ√ጉ	r Ası			o Gly	y Sei	r Met	Ala
136	Glı	ı GI	ı Ası	J GTZ			, IA.	L FII	y.	49	0				495	5
137			_		485	) . 7 ar	. 7 CI	n G11	. Ge	r Tri	o Tvi	r Tv:	r Lei	ı Ası	a Sei	Asn
138	Th:	r Gly	Tr			1 ASI	ı Ası	ıı Gı	50 50	, 5	- 2 -	1		510	0	
139				500	) 		- The	n T.e.	. Gl	n As	n Ası	n G1	v Se:	r Tri	р Ту	r Tyr
140	Gl	y Ala			ıını	( GT)	A TT	р <u>Бе</u> 52	ω <u>σ</u> ⊥.	10.		<b></b> .	52	5	-	_
141		_	519		_ ~1-		~ M^	- λ1 - λ1	o a Th	r ሬገ	v Tri	o Le			n Ası	n Gly
142	Le			a ASI	I GT	y se	53	5 A1			,,	54	0			
143	_	53	υ <u> </u>			, 7.~	23 17 n	⊃ Ne.	n Gl	v Se	r Me			r Gl	y Tr	p Leu
144	Se	r Tr	р Ту	r Ty:	г це	u AS	II AT	a Ab	11 GT	1 50						_

DATE: 07/07/1999 TIME: 14:28:17 PAGE: 4 RAW SEQUENCE LISTING

PATENT APPLICATION US/09/056,019A

Input Set: I056019A.RAW

145		545					550					555					560
146		Gln	Tyr	Asn	Gly	Ser	Trp	Tyr	Tyr	Leu	Asn	Ala	Asn	Gly	Ser	Met	Ala
147			•		•	565	-	-	-		570			_		575	
148		Thr	Gly	Trp	Leu	Gln	Tyr	Asn	Gly	Ser	Trp	Tyr	Tyr	Leu	Asn	Ala	Asn
149			•	-	580		-		-	585	-	-	-		590		
150		Gly	Asp	Met	Ala	Thr	Gly	Trp	Val	Lys	Asp	Gly	Asp	Thr	Trp	Tyr	Tyr
151		•	•	595			-	-	600	-	_	_	_	605			
152		Leu	Glu	Ala	Ser	Gly	Ala	Met	Lys	Ala	Ser	Gln	Trp	Phe	Lys	Val	Ser
153			610			•		615	•				620		-		
154		Asp	Lvs	Trp	Tyr	Tyr	Val	Asn	Gly	Ser	Gly	Ala	Leu	Ala	Val	Asn	Thr
155		625	-	-	-	-	630		_		_	635					640
156		Thr	Val	qaA	Gly	Tyr	Gly	Val	Asn	Ala	Asn	Gly	Glu	Trp	Val	Asn	
157				-	-	645	_				650			_		655	
158	<210>	SEQ	ID I	NO 3													
159	<211>																
160	<212>	TYPE	3: PI	RТ													
161	<213>				rept	toco	ccus	pnet	ımon:	iae							
162	<400>	SEQU	JENCI	E: 3	_			_									
163		Glu	Asn	Glu	Gly	Ala	Thr	Gln	Val	Pro	Thr	Ser	Ser	Asn	Arg	Ala	Asn
164		1				5					10					15	
165		Glu	Ser	Gln	Ala	Glu	Gln	Gly	Glu	${\tt Gln}$	Pro	Lys	Lys	Leu	Asp	Ser	Glu
166					20					25					30		
167		Arg	Asp	Lys	Ala	Arg	Lys	Glu	Val	$\operatorname{Glu}$	Glu	Tyr	Val	Lys	Lys	Ile	Val
168				35					40					45			
169		Gly	Glu	Ser	Tyr	Ala	Lys	Ser	Thr	Lys	Lys	Arg	His	Thr	Ile	Thr	Val
170			50					55					60				
171		Ala	Leu	Val	Asn	Glu	Leu	Asn	Asn	Ile	Lys	Asn	Glu	Tyr	Leu	Asn	Lys
172		65					70					75					80
173		Ile	Val	Glu	Ser	Thr	Ser	Glu	Ser	Gln	Leu	Gln	Ile	Leu	Met	Met	Glu
174						85					90					95	
175		Ser	Arg	Ser	Lys	Val	Asp	Glu	Ala		Ser	Lys	Phe	Glu	-	Asp	Ser
176					100					105	_				110		
177		Ser	Ser		Ser	Ser	Ser	Asp		Ser	Thr	Lys	Pro		Ala	Ser	Asp
178			_	115					120	_				125			
179		Thr		Lys	Pro	Asn	Lys		Thr	GLu	Pro	GTA		Lys	Val	Ala	GLU
180			130	_	_		~ 7	135		~ 7			140	<b>.</b>	•	<b>~1</b>	7 m
181			гÀг	гÀг	ьys	vaı	Glu	GIU	Ата	GIU	гуѕ	_	Ala	пуѕ	Asp	GIII	
182		145	<b>a</b> 1	3	<b>3</b>	3	150	m	D	Mile se	<b>T</b> 1.	155	III'u essa	T	mb .c	T 011	160
183		GIU	GIU	Asp	Arg		Asn	TYL	Pro	THE		THE	TAT	гуу	TILL		Gru
184		T 0	a1.,	T] ^	77-	165	000	7.00	171	C1.,	170	Ture	Tara	71-	C111	175	Clu
185		пеп	GIU	TTG	180	GIU	Ser	Asp	vaı	185	vaı	пуъ	цуъ	Ala	190	пеп	Gru
186 187		T 011	17n ]	Tara		Tarc	Ala	Aan	Clu		λνα	λen	Glu	Gln		Tla	Larg
188		Leu	val	195	v CL.L	пyэ	AT C	47971	200	110	J	עַטגי		205	د برد		_, 0
189		Gln	71 a		A 1 =	Glu	Val	Glu		Luc	Gln	Δla	Glu		Thr	Δrα	Len
190		J.111	210	OLU	WT.	Gru	VUL	215	DUL	-10			220			• 3	
191		Lve		Tle	Ive	Thr	Asp		Glu	Glu	Δla	Glu		G] 11	Ala	Lvs	Ara
192		225	_, 5		-10	~	230	3				235				-1-	240
193			Ala	Asp	Ala	Lvs	Glu	Gln	Glv	Lvs	Pro		Glv	Ara	Ala	Lys	-
194		3				245			1	-1-	250	1-				255	
						~											

DATE: 07/07/1999 RAW SEQUENCE LISTING TIME: 14:28:17 5 PAGE:

PATENT APPLICATION US/09/056,019A

Input Set: I056019A.RAW

```
Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys Glu Asn Asp Ala
195
                                             265
                        260
196
            Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu
197
                                         280
198
      <210> SEQ ID NO 4
199
      <211> LENGTH: 106
200
      <212> TYPE: PRT
201
      <213> ORGANISM: Streptococcus pneumoniae
202
      <400> SEQUENCE: 4
            Lys Pro Glu Lys Lys Val Ala Glu Ala Glu Lys Lys Val Glu Glu Ala
203
204
                                                  10
            Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr Pro
205
206
            Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp Val
207
208
            Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Glu Glu Ala Lys Glu
209
210
                                      55
            Pro Arg Asn Glu Glu Lys Val Lys Gln Ala Lys Ala Glu Val Glu Ser
211
212
                                  70
            Lys Lys Ala Glu Ala Thr Arg Leu Glu Lys Ile Lys Thr Asp Arg Lys
213
214
                                                   90
                              85
215
             Lys Ala Glu Glu Ala Lys Arg Lys Ala
216
                         100
217
       <210> SEQ ID NO 5
218
       <211> LENGTH: 109
 219
       <212> TYPE: PRT
 220
       <213> ORGANISM: Streptococcus pneumoniae
 221
       <400> SEQUENCE: 5
             Thr Glu Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val Glu Glu
 222
 223
             Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp Arg Arg Asn Tyr
 224
 225
                                               25
             Pro Thr Ile Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu Ser Asp
 226
 227
                                           40
             Val Glu Val Lys Lys Ala Glu Leu Glu Leu Val Lys Val Lys Ala Asn
 228
 229
                                       55
             Glu Pro Arg Asp Glu Gln Lys Ile Lys Gln Ala Glu Ala Glu Val Glu
 230
 231
                                                        75
             Ser Lys Gln Ala Glu Ala Thr Arg Leu Lys Lys Ile Lys Thr Asp Arg
 232
 233
                                                                        95
 234
             Glu Glu Ala Glu Glu Ala Lys Arg Arg Ala Asp Ala
 235
                          100
 236
       <210> SEQ ID NO 6
 237
        <211> LENGTH: 4
 238
       <212> TYPE: PRT
 239
       <213> ORGANISM: Streptococcus pneumoniae
 240
        <220> FEATURE:
  241
        <221> NAME/KEY: NON CONS
  242
        <222> LOCATION: (2)..(3)
        <223> OTHER INFORMATION: They could be any amino acid at these two
  243
```

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to

<223> fields of each sequence which presents at least one n or Xaa.

PAGE: 6

VERIFICATION SUMMARY

PATENT APPLICATION US/09/056,019A

DATE: 07/07/1999
TIME: 14:28:17

Input Set: 1056019A.RAW

Line	?	Error/Warning	Original Text								
797	W	"N" or "Xaa" used: Feature required "N" or "Xaa" used: Feature required "N" or "Xaa" used: Feature required	Lys Xaa Xaa Glu Xaa Glu Asn Glu Gly Glu Val Xaa Gly Glu Leu Ala Thr Pro Asp L								